

Sparse Coding Approach for Multi-Frame Image Super Resolution

Toshiyuki Kato¹, Hideitsu Hino², and Noboru Murata¹

¹Waseda University, 3-4-1 Ohkubo, Shinjuku, Tokyo, Japan,

²University of Tsukuba.

1-1-1 Tennodai, Tsukuba, Ibaraki, 305-8573, Japan

Keywords: image super-resolution, multi-frame super-resolution, sparse coding

Abstract

An image super-resolution method from multiple observation of low-resolution images is proposed. The method is based on sub-pixel accuracy block matching for estimating relative displacements of observed images, and sparse signal representation for estimating the corresponding high-resolution image. Relative displacements of small patches of observed low-resolution images are accurately estimated by a computationally efficient block matching method. Since the estimated displacements are also regarded as a warping component of image degradation process, the matching results are directly utilized to generate low-resolution dictionary for sparse image representation. The matching scores of the block matching are used to select a subset of low-resolution patches for reconstructing a high-resolution patch, that is, an adaptive selection of informative low-resolution images is realized. When there is only one low-resolution image, the proposed method works as a single-frame super-resolution method. The proposed method is shown to perform comparable or superior to conventional single- and multi-frame super-resolution methods through experiments using various real-world datasets.

1 Introduction

Super-resolution (SR) have been receiving a large amount of attention for creating clear images from low-resolution images. In stark contrast to simple picture interpolation techniques, SR methods utilize prior knowledges or assumptions on the structure and relationship of high- and low-resolution images. Development and spread of video equipments drive a growing need for SR techniques, which enable us to create high-resolution (HR) images from low-resolution (LR) images. See, e.g., [1, 2] for comprehensive surveys.

Super-resolution techniques can be divided into two categories, reconstruction-based SR [3–5], and example-based SR [6–9]. In the former approach, we compute HR images by simulating the image formation process. It is often used for multi-frame SR, where multiple LR images are used to obtain an HR image. A random Markov field model is usually adopted to represent the relationship between pixels of LR and HR images. This approach is intuitive and natural for SR. It is shown to give favorable results with an appropriate prior [10], but it often requires vast amounts of computation.

Example-based SR aims at inferring HR images based on small image segments extracted from training HR images. In many cases, it is adopted when we can use only one LR image. One of the representative works of example-based SR is the method by Freeman et al. [6], which is based on Neighborhood Embedding [11]. Recently, Yang et al [12, 13] improved [6] and proposed an SR method based on sparse coding [23]. Sparse coding is a methodology to represent observed signals with combinations of only a small number of basis vectors chosen from a large number of candidates. A set of basis vectors is called *dictionary*. A lot of single frame SR methods based on sparse coding are proposed, and they are experimentally shown to offer favorable HR images. In SR based on sparse coding, a dictionary for HR images is prepared in advance, and taking into account the image degradation process, a dictionary for LR images is generated from the HR dictionary. An LR image is represented by a combination of LR bases, then the coefficients for LR bases are used for combination coefficients of HR bases to obtain the sharpened image. In [12, 13], sparse coding is ingeniously utilized to obtain an HR image from only one LR observation. However, it should be possible to gain further improvements when we have multiple LR images for reconstructing an HR image.

In this paper, we propose a multi-frame SR method based on sparse coding. In general, multi-frame SR methods require registration of LR images, that is, we have to estimate relative displacements of LR images. We propose to use a sub-pixel accuracy block matching method for image registration. One of the contributions of the proposed method is in using the results of block matching to generate LR dictionaries. Another contribution is that the proposed method can adaptively select informative LR images for constructing HR image, which is also a beneficial side-effect of sub-pixel accuracy block matching. In that sense, the proposed method is *adaptively* selects the patches to be used for SR. By thresholding the matching score, the number of LR images used for reconstructing HR image varies in each small patch. We note that when there is only one LR image, the proposed method is the same as the conventional SR method based on sparse coding.

The rest of this paper is organized as follows. The image observation model is introduced in section II, and the sub-pixel accuracy block matching method is briefly explained without technical details in section III. The notion of sparse coding is introduced in section IV. In section V, conventional super-resolution methods based on sparse coding is explained, and in section VI, a novel multi-frame SR method based on sparse coding is proposed. Section VII shows experimental results, and the last section is devoted to concluding remarks.

2 Image Observation Model

In this section, we describe the image observation model. Following the idea of [4], we assume a continuous image $\tilde{X}(x, y)$ where $(x, y) \in \mathbb{R}^2$ are coordinate values. Then, we assume that an ideal discrete HR image \mathbf{X} and an LR image \mathbf{Y} are sampled from the continuous image \tilde{X} according to the following model:

$$X[m, n] = \left[\mathcal{W} \left(\tilde{X}(x, y) \right) \right] \downarrow_X \quad (1)$$

$$Y[m, n] = \left[\mathcal{H} * \mathcal{W} \left(\tilde{X}(x, y) \right) \right] \downarrow_Y + \mathcal{E}[m, n], \quad (2)$$

where \mathcal{W} and \mathcal{H} are warp and blur operators, \downarrow_X and \downarrow_Y are quantization operators to generate HR and LR images, and \mathcal{E} is an additive noise. In this paper, we denote coordinates in continuous space and discrete space by (x, y) and $[m, n]$, respectively. The blurring is expressed by the convolution operator $*$.

Following the conventional formulation of super resolution, we treat the HR and LR images as vectors $\mathbf{X} \in \mathbb{R}^{p_h}$, $\mathbf{Y} \in \mathbb{R}^{p_l}$, and the LR observation is assumed to be related with the HR image by

$$\mathbf{Y} = SHW\mathbf{X} + \boldsymbol{\varepsilon}, \quad (3)$$

where \mathbf{Y} and \mathbf{X} are column vectors with stacked images, the matrix $W \in \mathbb{R}^{p_h \times p_h}$ encodes the warping or spacial distortion, the matrix $H \in \mathbb{R}^{p_h \times p_h}$ models the blurring effect, $S \in \mathbb{R}^{p_l \times p_h}$ is the down-sampling operator, and $\boldsymbol{\varepsilon} \in \mathbb{R}^{p_l}$ is the Gaussian noise term.

Example-based SR approaches usually extract small patches $\mathbf{x} \in \mathbb{R}^{q_h}$ from the HR image and $\mathbf{y} \in \mathbb{R}^{q_l}$ from the LR image \mathbf{Y} . The whole HR image \mathbf{X} is obtained by integrating HR patches \mathbf{x} . With some abuse of notation of the operators and the noise term, each patch pair (\mathbf{x}, \mathbf{y}) is connected by the observation model

$$\mathbf{y} = SHW\mathbf{x} + \boldsymbol{\varepsilon}, \quad (4)$$

where $W \in \mathbb{R}^{q_h \times q_h}$, $H \in \mathbb{R}^{q_h \times q_h}$, $S \in \mathbb{R}^{q_l \times q_h}$, and $\boldsymbol{\varepsilon} \in \mathbb{R}^{q_l}$. The blurring effects are often modeled by convolution with a point spread function. In this paper, we use 9×9 Gaussian filter for the point spread function. The down-sampling process S is assumed to be an impulse sampling. The down-sampled image is further affected by the sensor noise and color filtering noise, and they are represented by an additive Gaussian noise term. In general, the spacial distortion W includes translation, rotation, deformation and other possible distortions. In this paper, we restrict W to simple translations, which is reasonable when we treat small patches instead of the whole image. In multi-frame SR, we assume that a set of N observed LR images $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_N$ are given. We first choose a *target* LR image \mathbf{Y}_t out of N observed images, and the final output of an SR method is the HR version of \mathbf{Y}_t . We refer other LR images as *auxiliary* LR images henceforth. Without loss of generality, we let $t = 1$, i.e., \mathbf{Y}_1 is the target image and $\mathbf{Y}_j, j = 2, \dots, N$ are the auxiliary images. The LR observations are related with the HR image \mathbf{X} by

$$\mathbf{Y}_j = S_j H_j W_j \mathbf{X} + \boldsymbol{\varepsilon}_j, \quad j = 1, \dots, N, \quad (5)$$

where $W_j \in \mathbb{R}^{p_h \times p_h}$ encodes relative displacement of a auxiliary image \mathbf{Y}_j to the target image \mathbf{Y}_1 in the HR image coordinate, and W_1 is the identity operator. The operator $H_j \in \mathbb{R}^{p_h \times p_h}$ models the blurring effect, $S_j \in \mathbb{R}^{p_l \times p_h}$ models the down-sampling, and $\boldsymbol{\varepsilon}_j \in \mathbb{R}^{p_l}$ is the Gaussian noise term.

3 Registration by Block Matching

In multi-frame SR, a reliable motion estimation is essential for achieving high-quality HR image. Image registration means not only identifying relative displacements between observed LR images, but also means the estimation of translation operators W_1, \dots, W_N in the model (5). Various motion estimation methods have been developed. In this paper, we adopt the block matching method [14, 15] as shown in Fig. 1, because the sparse coding method also extracts small patches from images. As mentioned earlier, we assume that positional relationship of blocks are fully described by parallel translation. Multi-frame SR techniques achieve remarkable results in resolution enhancement by estimating the sub-pixel shifts between images. There are a variety of strategies for estimating positional relationship in sub-pixel accuracy [16–21]. In this paper, we adopt the two-parameter simultaneous estimation method for similarity interpolation [22], which balances accuracy and computational efficiency. By the block matching

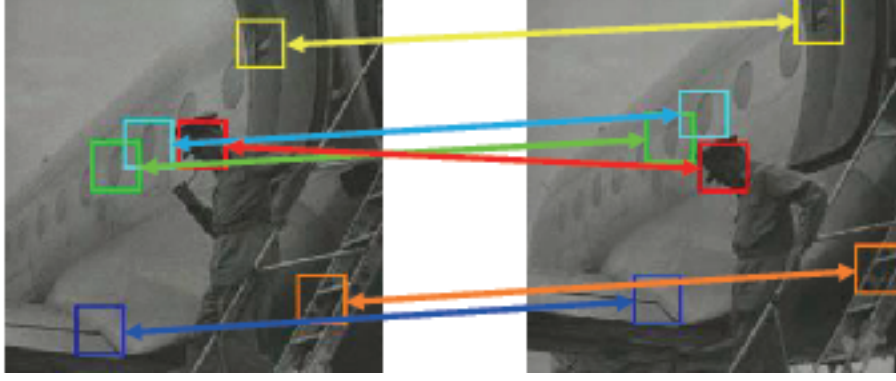


Figure 1: Conceptual figure of block matching

method, we obtain the similarity score

$$\text{sim}(\mathbf{y}_t, \mathbf{y}_j) = \frac{\mathbf{y}_t^\top \mathbf{y}_j}{\|\mathbf{y}_t\|_2 \cdot \|\mathbf{y}_j\|_2} \in [0, 1] \quad (6)$$

between patches from the target image \mathbf{Y}_1 and patches from auxiliary images \mathbf{Y}_j . Later, these scores are used to judge whether the patches extracted from auxiliary images should be used for the HR image reconstruction or not.

4 Sparse Coding

Sparse coding [23–26] is a methodology to represent observed signals with combinations of only a small number of basis vectors chosen from a large number of candidates. These basis vectors will be called *atoms* henceforth.

Let $\mathbf{D} = [\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_K] \in \mathbb{R}^{p \times K}$ be a dictionary which consists of K different atoms, $\boldsymbol{\alpha} \in \mathbb{R}^K$ be the coefficient vector for sparse representation of the signal $\mathbf{s} \in \mathbb{R}^p$. Then, the ℓ^0 -norm sparse coding problem is formally written as

$$\text{minimize } \|\boldsymbol{\alpha}\|_0 \quad \text{subject to} \quad \|\mathbf{D}\boldsymbol{\alpha} - \mathbf{s}\|_2^2 \leq \epsilon, \quad (7)$$

where ϵ is a predefined tolerance for approximation, or

$$\text{minimize } \|\mathbf{D}\boldsymbol{\alpha} - \mathbf{s}\|_2^2 \quad \text{subject to} \quad \|\boldsymbol{\alpha}\|_0 \leq T, \quad (8)$$

where T is a predefined number of atoms to be used for representation. The ℓ^0 -norm $\|\boldsymbol{\alpha}\|_0$ is defined by the number of nonzero elements of a vector $\boldsymbol{\alpha}$

$$\|\boldsymbol{\alpha}\|_0 = \#\{k | \alpha_k \neq 0, k \in \{1, \dots, K\}\}. \quad (9)$$

Given a dictionary \mathbf{D} , the problems (7) and (8) are known to be *NP-hard* [27], and the matching pursuit type algorithms [28–30] are often used for obtaining approximate solutions. As another approach for sparse coding, the problem is relaxed to an ℓ^1 -norm minimization problem as

$$\text{minimize } \|\boldsymbol{\alpha}\|_1 \quad \text{subject to} \quad \|\mathbf{D}\boldsymbol{\alpha} - \mathbf{s}\|_2^2 \leq \epsilon. \quad (10)$$

This problem (10) adopts the ℓ^1 -norm of coefficients as a measure of sparsity, and referred to as the ℓ^1 -norm sparse coding. Using the Lagrange multiplier η , this problem is rewritten as

$$\text{minimize } \frac{1}{2} \|\mathbf{D}\boldsymbol{\alpha} - \mathbf{y}\|_2^2 + \eta \|\boldsymbol{\alpha}\|_1 \quad (11)$$

where $\eta \geq 0$ represents the strength of the sparseness constraint. In [31], it is argued that the ℓ^1 -norm sparse coding is more stable than the ℓ^0 -norm sparse coding. The problem (11) is the same form as Lasso [32], which is widely used sparse linear regression, and there are some computationally efficient algorithms for estimating the coefficient such as LARS-lasso algorithm [33] and feature-sign search algorithm [36]. In this paper, we adopt the ℓ^1 -norm sparse coding as a building block for super-resolution, and use the LARS-lasso algorithm for solving the problem (11).

In sparse coding, the design of dictionary \mathbf{D} is of prime importance. The first attempt of the example-based SR method employs small patches of HR images themselves as atoms [6], though, learning the dictionary from observed data is shown to improve the reconstruction performance. Actually, the HR and LR dictionary learning is the core technical component of the state-of-the-art SR methods based on sparse coding [12, 13, 41]. Representative methods for dictionary learning are Method of Optimal Directions [34] which is based on vector quantization, and K-SVD [35] which is based on k-means clustering and the singular value decomposition. In this paper, we use Lee's method [36], which is known to perform well with reasonable computational cost.

5 Super-Resolution via Sparse Coding

The basic idea of single-frame SR based on sparse coding is proposed in [6]. We first explain that idea and techniques for improving the quality of the reconstructed image. Then, we introduce the multi-frame SR method.

5.1 Inferring HR Image From Single LR image

Using HR training images and LR training images, an HR dictionary $\mathbf{D}_h \in \mathbb{R}^{q_h \times K}$ and an LR dictionary $\mathbf{D}_l \in \mathbb{R}^{q_l \times K}$ are learned before performing HR image reconstruction. Assume the HR patch \mathbf{x} is represented by $\mathbf{x} = \mathbf{D}_h \boldsymbol{\alpha}$, and represent the image degradation process SHW by $L \in \mathbb{R}^{q_l \times q_h}$. Then, the LR and HR images are connected as

$$\mathbf{y} = L\mathbf{x} = L\mathbf{D}_h \boldsymbol{\alpha} = \mathbf{D}_l \boldsymbol{\alpha}, \quad (12)$$

where $\mathbf{D}_l = SHW\mathbf{D}_h$. In the above equation, we made an assumption, which is the core idea of sparse coding based SR, that the HR patch shares the same coefficient with the LR patch in sparse representation. For each input LR patch \mathbf{y} , a sparse representation with respect to \mathbf{D}_l is found. Then, the corresponding HR bases in \mathbf{D}_h is combined according to these coefficients to generate the output HR patch \mathbf{x} (Fig. 2).

5.2 Reducing Artifact at Patch Boundaries

In patch-based approaches, a sparse representation is obtained first for an LR patch then for the corresponding HR patch. Obtaining sparse representation individually for each local patch does not guarantee the compatibility between adjacent patches, and we often suffer from block artifacts at the patch boundaries. To reduce these block artifacts, overlapping patches are often used. In [6] and [12], an additional optimization problem is solved after patch-wise SR to

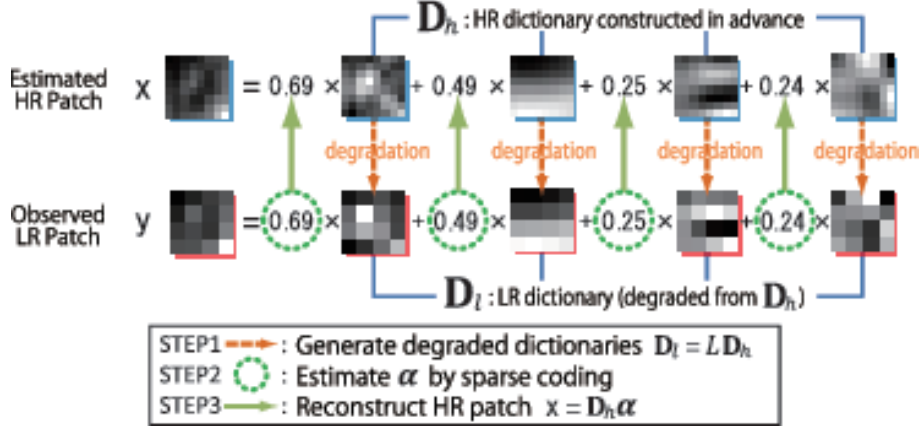


Figure 2: A framework of super resolution based on sparse coding.

improve the consistency of adjacent patches. In our proposed method, for the sake of computational efficiency, we adopt a simple weighted averaging method with 2-dimensional hanning window for combining adjacent HR patches.

5.3 Maintaining Global Consistency by Back-Projection

Sparse coding of the form Eq. (11) does not demand exact equality $D\alpha = y$. Because of this and also because of noise and patch-wise processing, the fundamental assumption denoted by Eq. (3) for the target image Y_t and X may not be satisfied after patch-wise SR. Hence, the back-projection [37] is performed for maintaining the global consistency [12, 13, 39]. Let X_0 be the obtained high-resolution image. To minimize the discrepancy between X_0 and the HR image under the model (3), we solve the following minimization problem

$$\underset{X}{\text{minimize}} \|SHX - Y\|_2^2 + c\|X - X_0\|_2^2 \quad (13)$$

by gradient descent, where $c > 0$ is a trade-off parameter. The updating formula of the gradient method is explicitly written by

$$X_{t+1} = X_t - \nu[H^T S^T (SHX_t - Y) + c(X - X_0)], \quad (14)$$

where $\nu > 0$ is a step size parameter.

5.4 Multi-frame Super Resolution by Sparse Coding

In multi-frame scenario, each LR image is assumed to be an outcome of different degradation of the same HR image. Multi-frame SR is expected to achieve better results using multiple LR images, though, extraction and integration of useful information in LR images are not a trivial task.

In images observation model (5), we assume that $H_j = H$ and $S_j = S$, $\forall j \in \{1, \dots, N\}$ in Eq. (5) for the sake of simplicity. We also assume ε_j for all j follows the same Gaussian distribution. From observed LR images $Y_j, j = 1, \dots, N$, patches $y_j, j = 1, \dots, N$ with different sizes are extracted by clipping operators $C_j \in \mathbb{R}^{q_{lj} \times p_{lj}}$, $j = 1, \dots, N$, which will be defined later.

As is the case with single-frame SR, let \mathbf{X} denote the HR image to be reconstructed, \mathbf{x} be a patch extracted from \mathbf{X} , $\mathbf{Y}_1, \dots, \mathbf{Y}_N$ be LR images and $\mathbf{y}_1, \dots, \mathbf{y}_N$ be patches of LR images corresponding to \mathbf{x} . For realizing multi-frame SR, we solve the following optimization problem:

$$\underset{\alpha}{\text{minimize}} \|\tilde{\mathbf{D}}_l \alpha - \tilde{\mathbf{y}}\|_2^2 + \eta \|\alpha\|_1 \quad (15)$$

$$\text{with } \tilde{\mathbf{D}}_l = \begin{bmatrix} \mathbf{D}_{l1} \\ \vdots \\ \mathbf{D}_{lN} \end{bmatrix} \quad \text{and} \quad \tilde{\mathbf{y}} = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_N \end{bmatrix}. \quad (16)$$

In our framework of multi-frame SR, we first select a target image from a set of observed LR images, and construct the HR image of the target image using other auxiliary LR images. Then, \mathbf{y}_1 and \mathbf{D}_{l1} can be extracted and constructed in the same manner as single-frame SR. We have to extract and construct other $N - 1$ patches and dictionaries.

In [38, 39], to alleviate the inaccuracy of integer-pixel accuracy block matching, referring the technique in the literature of noise-reduction [40], they used similarities calculated in the block matching procedure as weights for integrating the coefficients obtained by sparse coding of each auxiliary patches \mathbf{y}_j , $j = 2, \dots, N$. Their method does not involve explicit sub-pixel accuracy block matching. By adopting sub-pixel accuracy block matching, we can expect higher accuracy in the HR image reconstruction.

6 Proposed Method

An important characteristic of the proposed multi-frame super resolution method based on sparse coding is in considering the sub-pixel accuracy image registration. In the following three subsections, we explain the basic idea and the concrete procedure for the proposed method¹.

6.1 General Description

We propose to generate a stacked dictionary $\tilde{\mathbf{D}}_l \in \mathbb{R}^{q_J \times K}$ and stacked observation vectors $\tilde{\mathbf{y}} \in \mathbb{R}^{q_J}$, taking into account the displacements between LR images in sub-pixel level. To encourage the intuitive understanding, in Fig. 3, we show a conceptual diagram in which images are replaced with one-dimensional signals. In this figure, horizontal axis and vertical axis represent location in the image and luminance value, respectively.

In Fig. 3, bottom three images (g), (h), and (i) correspond to the given N observed LR images, and their HR counterparts are shown as (d), (e), and (f). We suppose these LR and HR images are sampled from continuous images (a), (b), and (c), that is, Eqs. (1) and (2) show relationships between (a) and (d), and (a) and (g), respectively. On the other hand, Eq. (3) shows relationship between (d) and (g). Given the shift information, continuous images (a), (b), and (c) are inter-commutable. On the other hand, because of discretization, it is not possible to convert one discrete HR image to another precisely. In Fig. 3, solid arrows depict the image degradation processes, dotted arrows depict atom generation processes, and dashed lines depict positional relationships in sub-pixel accuracy.

Observed LR images $\mathbf{Y}_1, \dots, \mathbf{Y}_N$ are assumed to be down-sampled and blurred versions of continuous images, and they share the same degradation process except warping process. Let patches $\mathbf{y}_1, \dots, \mathbf{y}_N$ be obtained by clipping LR images $\mathbf{Y}_1, \dots, \mathbf{Y}_N$ taking into account the sub-pixel accuracy shifts estimated by the block matching procedure introduced in section 3.

¹A simple implementation of the proposed method will be available from author's website or journal's support website.

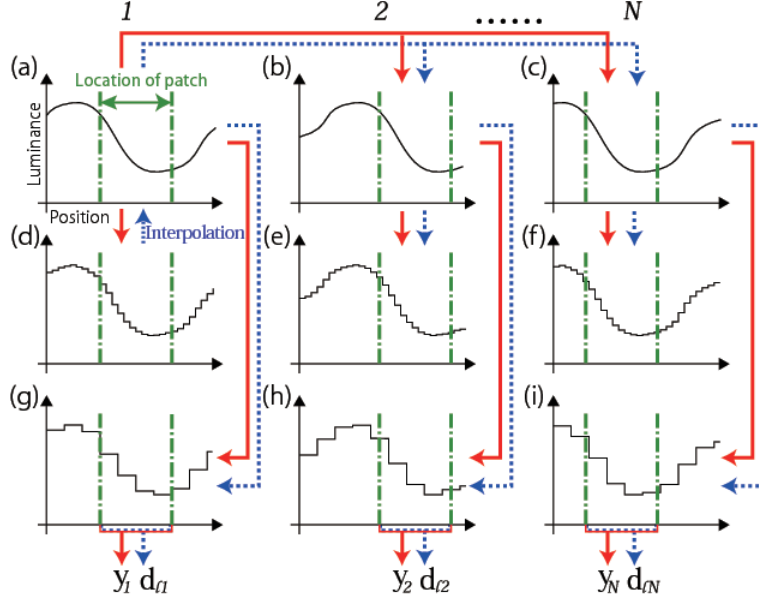


Figure 3: Sampling considering sub-pixel shift.

After clipping LR patches $\mathbf{y}_1, \dots, \mathbf{y}_N$, we generate LR atoms from the HR atoms learned by a dictionary learning algorithm. Suppose the signal shown in Fig. 3 (d) is one of the HR atoms. We can approximate the continuous atom (a) by interpolation of the atom in (d), and, by parallel shift for continuous images, we can also approximate warped continuous atoms (b) and (c). Then, we obtain warped LR atoms (h) and (i) from (b) and (c). In actual implementation, as described in the next section, we will calculate warped HR atoms (e) and (f) directly from (d) by interpolation, and obtain (h) and (i) by applying blur and down-sampling operators.

6.2 Implementation

In this section, we present a concrete procedure for generating the stacked observation vector $\tilde{\mathbf{y}}$ and stacked LR dictionary $\tilde{\mathbf{D}}_l$, which are the core components of the proposed SR method.

For generating the stacked observation vector, patches $\mathbf{y}_1, \dots, \mathbf{y}_N$ are extracted from auxiliary LR images by clipping operators $C_j \in \mathbb{R}^{q_{lj} \times p_l}$ $j = 1, \dots, N$. The clipping operators are designed to extract only the pixels inside the cutoff line as shown in Fig. 4. Then, the clipped images are stacked to construct $\tilde{\mathbf{y}}$ as

$$\tilde{\mathbf{y}} = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_N \end{bmatrix} = \begin{bmatrix} C_1 \mathbf{Y}_1 \\ \vdots \\ C_N \mathbf{Y}_N \end{bmatrix}. \quad (17)$$

We next explain how to generate the stacked LR dictionary $\tilde{\mathbf{D}}_l$. As shown in Eq. (15), $\tilde{\mathbf{D}}_l$ is composed of dictionaries \mathbf{D}_{lj} , $j = 1, \dots, N$ for \mathbf{y}_j , $j = 1, \dots, N$. Each dictionary \mathbf{D}_{lj} is generated as follows. Remember that we already have a dictionary \mathbf{D}_h for HR images. Each HR atom in \mathbf{D}_h is embedded into the HR patch region, then degenerated according to the image degradation processes S and H to obtain LR atoms (see Fig. 3 (d) and (g), and Fig. 5: transition from HR space to LR space). Parallel translation in HR space is performed using bi-linear interpolation [46]. When the luminance value at (x, y) is estimated, we define $u_1 = \lfloor x \rfloor$ and

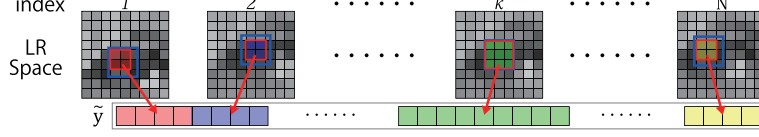


Figure 4: The procedure of generating stacked LR observation \tilde{y} by clipping operation

$u_2 = \lfloor y \rfloor, v_1 = x - u_1, v_2 = x - u_2$. Then, we obtain the luminance value at (x, y) as

$$X(x, y) = \begin{bmatrix} 1-v_1 & v_1 \end{bmatrix} \begin{bmatrix} X[u_1, u_2] & X[u_1+1, u_2] \\ X[u_1, u_2+1] & X[u_1+1, u_2+1] \end{bmatrix} \begin{bmatrix} 1-v_2 \\ v_2 \end{bmatrix}. \quad (18)$$

To obtain an operator W for this sub-pixel accuracy translation, we calculate the weighted sum of four pixel-level warp matrices. Let $w_1, w_2 \in \mathbb{N}$ be horizontal and vertical shifts, respectively, and $W_{[w_1, w_2]}$ be the warp operator. Then, the operator W is obtained by

$$W = (1-v_1)(1-v_2)W_{[w_1, w_2]} + (1-v_1)v_2W_{[w_1+1, w_2]} + v_1(1-v_2)W_{[w_1, w_2+1]} + v_1v_2W_{[w_1+1, w_2+1]}. \quad (19)$$

Atoms for LR images are calculated by clipping operation from the degraded HR atoms. Then, they are stacked to obtain $\tilde{\mathbf{d}}_l$ as shown in Fig. 5. From every atom in HR dictionary \mathbf{D}_h , we calculate the stacked LR atom $\tilde{\mathbf{d}}_l$ as presented above, and obtain the stacked LR dictionary $\tilde{\mathbf{D}}_l$. Denoting the embedding operator $R \in \mathbb{R}^{p_h \times q_h}$, which embeds an HR atom into the location of target HR patch, $\tilde{\mathbf{D}}_l$ is represented by

$$\tilde{\mathbf{D}}_l = \begin{bmatrix} \mathbf{D}_{l1} \\ \vdots \\ \mathbf{D}_{lN} \end{bmatrix} = \begin{bmatrix} C_1 S H W_1 R \mathbf{D}_h \\ \vdots \\ C_N S H W_N R \mathbf{D}_h \end{bmatrix}. \quad (20)$$

6.3 Summary of the Proposed Method

We summarize the proposed method as a concrete procedure in Algorithm 1. First of all, we arbitrary fix the target image from a set of N observed images. We let the first image as the target image for the sake of notational simplicity. The output of the algorithm is the HR version of this target image \mathbf{Y}_1 . The rest of the algorithm is composed of two stages, the super-resolution stage, and the post-processing stage. It is supposed that an HR dictionary is learned from a set of HR images beforehand. In the super-resolution stage, the following procedures are performed for every small patch \mathbf{y}_1 extracted from the target image \mathbf{Y}_1 , and then the tentative HR image \mathbf{X}_0 is obtained by weighted average of the small patches, as described in 5.2. By block matching method, relative displacements of LR patches extracted from LR images $\mathbf{Y}_j, j = 2, \dots, N$ are estimated in sub-pixel accuracy. Then, based on this registration results, the stacked image and stacked dictionary are generated as shown in Fig. 4 and Fig. 5, respectively. In an LR image $\mathbf{Y}_j, j = 2, \dots, N$, if there are no block with similarity higher than a predefined threshold value, we do not use other LR images than \mathbf{Y}_1 . Then, the coefficient α is estimated for sparse representation of $\tilde{\mathbf{y}}_l$ by the dictionary $\tilde{\mathbf{D}}_l$. We note that when estimating the coefficient α by sparse coding algorithm, the mean value m of $\tilde{\mathbf{y}}$ is removed from each element of $\tilde{\mathbf{y}}$, and added to each element of $\mathbf{D}_h \alpha$ when reconstructing HR patch \mathbf{x} . Finally, by the HR dictionary \mathbf{D}_h and coefficient α , the HR patch $\hat{\mathbf{x}}$ is constructed. In the post-processing stage, the collection of the tentative HR images \mathbf{X}_0 is revised by gradient method as stated in 5.3.

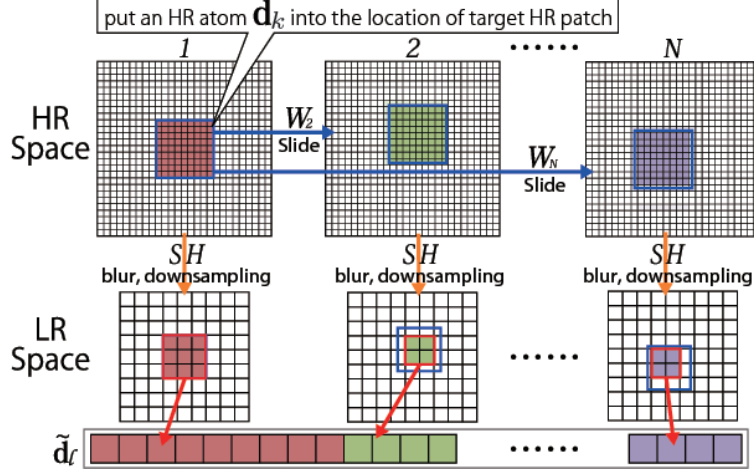


Figure 5: The procedure of generating stacked LR dictionary $\tilde{\mathbf{d}}_{lk}$.

6.4 Relationship to Other Methods

Finally, we conclude this section by comparing the proposed method to existing SR methods based on sparse coding, and elucidate the advantages of the proposed method. The comparison among various methods is summarized in Tab. 1.

To the best of the authors' knowledge, there are only one extant multi-frame SR method based on sparse coding [39]. In [39], all of the observed LR images are used for HR image construction, while in the proposed method, only informative LR images for reconstruction are selected for each patch.

Multi-frame SR requires accurate registration of observed LR images. In our proposed method, the block matching method with sub-pixel accuracy [22] is adopted, and the matching result is efficiently utilized for stacked LR dictionary construction. Using the similarity scores obtained in this registration step, we select informative patches to be used for HR image reconstruction.

In sparse coding stage of SR, both the ℓ^0 -norm and the ℓ^1 -norm are possible for the regularization of the coefficients. The proposed method adopt the ℓ^1 -norm, because we observed that it is more stable than the ℓ^0 -norm for SR tasks. We used the LARS-Lasso algorithm for efficiently solving the sparse coding problem (11).

For learning the HR dictionary \mathbf{D}_h , we apply Lee's method [36] which is known to be fast and accurate. The LR dictionary is generated by atoms in \mathbf{D}_h according to the assumed blur, down-sampling and estimated translation. This strategy of generating an LR dictionary is more faithful to the idea of super-resolution than the joint dictionary learning strategy adopted in [13, 39].

In order to focus on perceptually important high-frequency content of the image, Yang et al. [12] proposed to apply four differentiation filters to LR images and applied sparse coding for these extracted features. Feature extraction may improve the reconstruction accuracy, though, to keep the algorithm simple and to avoid increasing of computational costs, we did not use the feature extraction filter and obtained satisfiable results.

Algorithm 1 Multi-frame Super-Resolution via Sparse Representation

Input: LR Images $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_N$ and HR Dictionary \mathbf{D}_h . A threshold $\delta \in [0, 1]$.

{Super-Resolution Stage}

Select target image (\mathbf{Y}_1 for example)

for each patch \mathbf{y} **do**

-[Step 1] Perform block matching and select informative patches based on δ

-[Step 2] Make $\tilde{\mathbf{y}}, \tilde{\mathbf{D}}_l$ {cf. section 3 and Figs. 4, 5}

-[Step 3] Perform sparse coding

$m \leftarrow \text{mean}(\tilde{\mathbf{y}})$

$\tilde{\mathbf{y}} \leftarrow \tilde{\mathbf{y}} - m$

$\alpha \leftarrow \text{SparseCoding}(\tilde{\mathbf{D}}_l, \tilde{\mathbf{y}})$

-[Step 4] Reconstruction

$\mathbf{x} \leftarrow \mathbf{D}_h \alpha + m$

end for

Make \mathbf{X}_0 from all patches \mathbf{x} {cf. section 5.3}

{Post-Processing Stage}

repeat

$\mathbf{X} \leftarrow \mathbf{X} - \nu[H^T S^T(SH\mathbf{X} - \mathbf{Y}) + c(\mathbf{X} - \mathbf{X}_0)]$

until convergence

Table 1: Comparison to extant works on sparse coding based SR.

	input	registration	sparse coding	dictionary learning	feature extraction	Local consistency at patch boundary	post-processing
proposed	multiple	sub-pixel block-matching	ℓ^1 norm SC (LARS-Lasso)	\mathbf{D}_h : Lee's method \mathbf{D}_l : $L(\mathbf{D}_h)$	–	weighted averaging	back-projection
P. Wang et al. (2011) [39]	multiple	pixel block-matching	ℓ^1 norm SC	\mathbf{D}_h and \mathbf{D}_l : Lee's method (Joint Dictionary Learning)	differentiation filter	weighted SC	back-projection
J. Yang et al. (2008) [12]	single	–	ℓ^1 norm SC	\mathbf{D}_h : random sampling \mathbf{D}_l : $L(\mathbf{D}_h)$	differentiation filter	weighted SC	back-projection
J. Yang, et al. (2010) [13]	single	–	ℓ^1 norm SC	\mathbf{D}_h & \mathbf{D}_l : Lee's method (Joint Dictionary Learning)	differentiation filter	weighted SC	back-projection
R. Zeyde et al. (2010) [41]	single	–	ℓ^0 norm SC (OMP)	: find the best \mathbf{D}_h dictionary with \mathbf{D}_l \mathbf{D}_l : K-SVD	differentiation filter	averaging (overlap is taken account in dictionary learning)	–
L. Min et al. (2010) [43]	single	–	ℓ^1 norm SC	\mathbf{D}_h & \mathbf{D}_l : Sparse K-SVD [42] (Joint Dictionary Learning)	differentiation filter	weighted SC	back-projection
W. Dong et al. (2011) [44]	single	–	Iterated Shrinkage Algorithm for ASDS with AReg	Sub-dictionaries for ASDS: K-means and PCA	high-pass filter	averaging	–
X. Lu et al. (2012) [45]	single	–	Geometry Constrained ℓ^1 norm SC	: find the best \mathbf{D}_h dictionary with \mathbf{D}_l \mathbf{D}_l : Geometry Dictionary	–	averaging	–

7 Experimental Results

In this section, we first demonstrate the super-resolution results obtained by the proposed method on some sets of still images, and some sets of images from movies.

In our experiments, we magnify the input LR images by a factor of 3 for all cases. We used 5 LR images to estimate an HR images for multi-frame SR methods.

7.1 Application to Still Images

In the experiments with still images, we generated artificially degraded LR images according to the generative model in Eq. (5). As described in section II, we suppose observed LR images are generated from an HR image through parallel translations, blurring, down-sampling, and addition of noises. The parallel translations are imitated by shifting the HR image. The degree of vertical and horizontal shifts are randomly sampled from a uniform distribution in $[-5, 5]$. The blurring operation H is realized by convolution of 9×9 -pixel Gaussian filter with the standard deviation $\sigma_h = 1$. The blurred and shifted images are then down-sampled by the factor 3. Finally, noises sampled from $\mathcal{N}(0, \sigma_n = \sqrt{2})$ are added to generate LR observation images. We note that by this noise addition, SNRs of resulting 5 images are about $37 \sim 40$ [dB].

We randomly selected one target LR image, and estimated relative displacements from the target to other four images by sub-pixel accuracy block matching. In our experiments, both the intensity of the blur and the noise are supposed to be given. We summarize the experimental settings in table 2.

Table 2: Experimental settings (for still images)

parameters	values
number of LR images	5
parallel translation	Unif($-5, 5$) [px]
blur	9×9 Gaussian filter ($\sigma_h = 1$)
noise	Gaussian ($\sigma_n = \sqrt{2}$)

We compare the proposed method to four conventional methods. The first method is bi-cubic interpolation. This method is simple and regarded as a baseline for SR. The second method is the one proposed in [12], which is a single-frame SR method based on joint dictionary learning. We refer to this method as SF-JDL (Single Frame super resolution based on Joint Dictionary Learning). The third method is the one proposed in [41], which is based on the ℓ^0 sparse coding and we refer to this method as SF-L0 (Single Frame super resolution based on the ℓ^0 sparse coding). The fourth method is the one proposed in [39], which is a multi-frame SR method based on joint dictionary learning. We refer to this method as MF-JDL (Multi-Frame super resolution based on Joint Dictionary Learning). As for MF-JDL, we used 5 LR images including a target image for reconstructing HR images from them. As for bi-cubic method, SF-JDL and SF-L0, only the target LR image is used for HR image reconstruction. As for the relative displacements of LR images in the proposed methods, we tried two settings. The one is estimating the displacements by sub-pixel accuracy block matching method, and the other is assuming the ground-truth displacements are given. The former is referred to as proposed(1), while the later is referred to as proposed(2) in this section.

For quantitative comparison of SR methods, we use the Peak Signal to Noise Ratio (PSNR)

defined as

$$\text{PSNR[dB]} = 10 \log_{10} \frac{255^2}{\text{MSE}}, \quad (21)$$

where MSE is the mean squared error between the original HR image and the estimated HR image, and the higher PSNR indicates the better SR performance. We note that SF-L0 is based on the second order differential of images, and it can not reconstruct three pixels from the edge (rim) of the image. So, we calculate the PSNR discarding three pixels from the edge of the reconstructed HR images, and this evaluation may work advantageous to SF-L0.

There are seven parameters in the proposed method:

1. the strength of sparseness constraint η in sparse coding
2. the number of atoms m in the dictionary used in sparse coding
3. the size of patch q_h
4. the size of overlapping of patches
5. the threshold δ to judge whether a LR patch is used or not based on similarity score obtained by block matching
6. the balancing parameter c for global consistency (13)
7. the step size ν for back-projection (14)

According to our preliminary experimental results, we set the values of these parameters as shown in table 3. We note that the essentially new parameter in our method is δ . For other SR methods compared in this paper, other six parameters are also optimized according to our preliminary experiments. In this experiment, we set the threshold δ to zero and use all the 5 LR images for HR image reconstruction. We use two different gray-scale images (Lena and Cam-

Table 3: Parameter settings for the proposed method (still image)

parameters	values
η : penalty for the ℓ^1 -norm in (15)	0.05
m : number of basis	512
q_h : size of HR patch	$15 \times 15 = 125$
size of LR patch overlapping	3
δ : threshold of block matching score	0.0
c : trade-off in (14)	0.0001
ν : step size in (14)	0.001

eraman), and three color images (Flower, Girl, and Parthenon) for evaluating the performance of SR methods. When we deal with color images, we first convert the image into YCbCr format, then apply SR methods only to luminance channel (Y). Values of other channels Cb and Cr are simply expanded by bi-cubic interpolation.

We show the experimental results in Fig. 6 to Fig. 7. We note that we also applied SR methods to the famous color images (Flower, Girl, and Parthenon), but for the sake of image size, we omit showing the figures and just show the performance in PSNR measure. In these

figures, the degraded LR image, the results by bi-cubic interpolation, SF-JDL, SF-L0, MF-JDL, proposed(1), proposed(2), and the original HR image are shown. These figures indicate that the proposed method can generate much higher resolution images with sharp edges. Not surprisingly, the proposed(2), where the ground-truth displacements are given, provides the most clear HR images. However, proposed(1), where the displacements are estimated, provides clearer HR images than that obtained by other conventional methods.

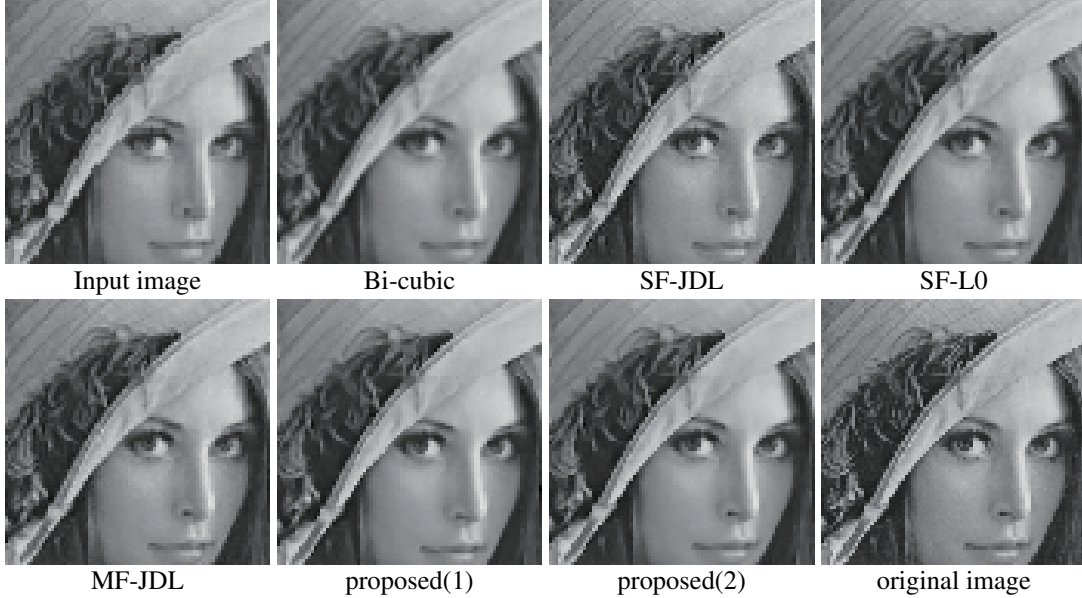


Figure 6: Images estimated from LR observations (Lena)

We also show PSNR values obtained by various methods in table 4. For evaluating the PSNR values, we randomly generated 100 sets of 5 warp operators and generated degraded images by adding random observation noises. From each set of observed images, we randomly choose one target image. The only target image is used for single-frame SR, while in multi-frame SR, the remaining 4 images are used as auxiliary LR images. The means and standard deviations of PSNR values are calculated using 100 SR results by each methods. The best and the second best results are shown in bold style. From table 4, the proposed method is shown to outperform

Table 4: PSNRs of SR methods

Image	Input	Bi-cubic	SF-JDL	SF-L0	MF-JDL	proposed(1)	proposed(2)
Lena	26.34 \pm 0.00	28.01 \pm 0.01	28.64 \pm 0.02	29.38 \pm 0.01	29.10 \pm 0.01	29.97 \pm 0.12	30.94 \pm 0.20
Camerman	24.37 \pm 0.00	26.87 \pm 0.01	27.85 \pm 0.02	29.08 \pm 0.01	28.17 \pm 0.02	29.72 \pm 0.41	32.24 \pm 0.35
Flower	33.44 \pm 0.00	35.36 \pm 0.01	35.52 \pm 0.02	36.36 \pm 0.01	36.00 \pm 0.02	36.44 \pm 0.11	36.92 \pm 0.15
Girl	29.81 \pm 0.00	31.06 \pm 0.00	31.33 \pm 0.01	31.74 \pm 0.01	31.55 \pm 0.01	31.88 \pm 0.06	32.20 \pm 0.10
Parthenon	23.50 \pm 0.00	24.31 \pm 0.00	24.41 \pm 0.00	24.92 \pm 0.00	24.59 \pm 0.00	25.36 \pm 0.10	26.39 \pm 0.15

other conventional methods. We note that standard deviations of PSNR values for all of the single-frame SR methods are very low, because randomness involved in these methods is only due to observation noise. The standard deviations of MF-JDL is also low compared to those of the proposed methods. The relatively high standard deviations for the proposed methods can be attributed to instability of the sub-pixel accuracy block matching method. However, even

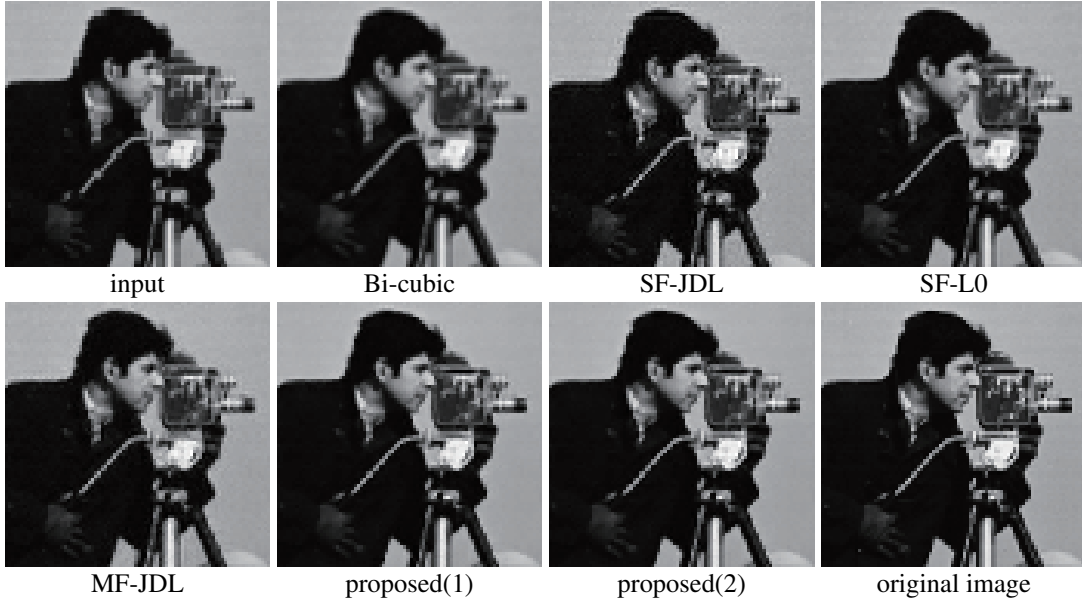


Figure 7: Images estimated from LR observations (Cameraman)

with relatively high standard deviations, the proposed methods show statistically significant improvements to conventional methods.

7.2 Application to Motion Pictures

To see that the proposed method can be applied to more practical problems, we perform experiments using LR images sequentially captured from movies. Images captured from movies are degraded according to the degradation model specified by table 2 excluding the parallel translation. From five consecutive LR images, the middle (third in the temporal sequence) image is selected as the target image, and other four are considered as auxiliary images.

Except for the threshold δ , the same parameters shown in table 3 are used in this section. The parameter δ determines whether an LR image is used for HR image reconstruction or not based on the matching score (6). This parameter is particularly important in multi-frame SR applied for LR images captured from movies. In this experiment, we show experimental results with $\delta = 0.003$ and $\delta = 0.001$. When the proposed method is applied to LR images sequentially captured from movies, an object appears in an LR image can be hidden behind another object. It is also possible that the object goes out of the scene. When the threshold is set to $\delta = 0.003$, only the auxiliary patches which obtain high similarity to the target patch are used for the HR reconstruction. On the other hand, when the threshold is set to $\delta = 0.001$, the algorithm adopts most of patches cropped from auxiliary LR images.

We use a gray-scale movie (MacArthur) and two different color movies (Samurai and Heading) for evaluating the performance of SR methods. We capture five consecutive images from each movie, and make LR images according to the degradation model. The obtained HR images using various SR methods are shown in Fig. 8, Fig. 9, and Fig. 10.

From Fig. 8, Fig. 9, and Fig. 10, the HR images obtained by the proposed method are clear and have distinct edges. By comparing to the original HR images, we can see that the proposed method can reconstruct edges and line-like features in the original images. On the other hand, there are patches in which the proposed method can not clearly reconstruct textures like wrinkles

and patterns in clothes. Also, in the case of “Heading”, the overall sharpness of the image is improved but undesirable artifacts are produced. This is possibly because the original movie contains abrupt shifts.

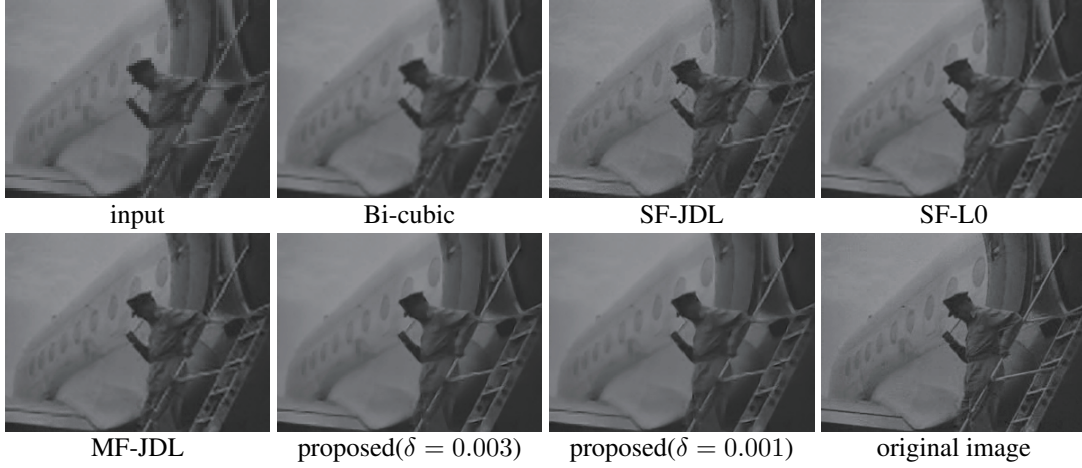


Figure 8: Images estimated from LR observations (MacArthur)

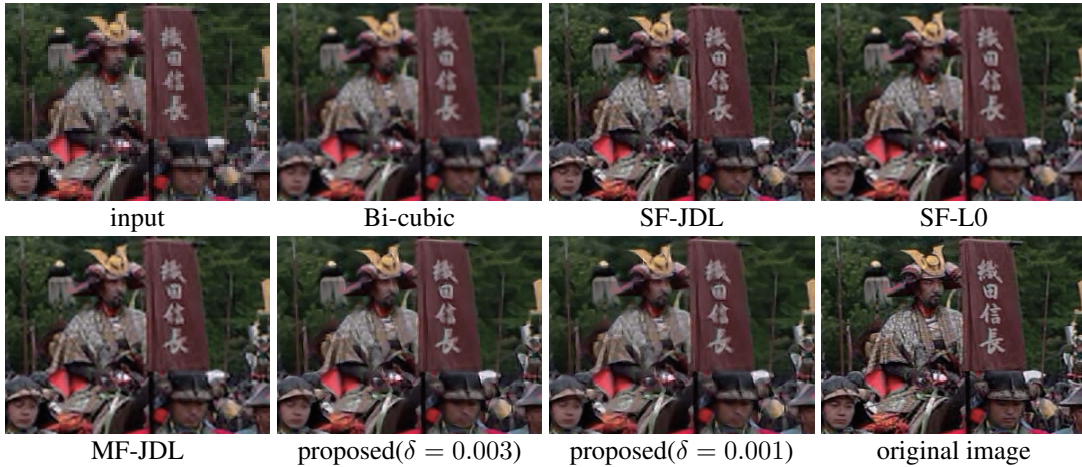


Figure 9: Images estimated from LR observations (Samurai)

8 Conclusion

In this paper, we proposed a multi-frame SR method based on sparse coding and sub-pixel accuracy block matching. The main contribution of the present paper is in proposing a natural extension of the single-frame SR method based on the sparse coding, with a novel combination of sub-pixel accuracy block matching method and an LR atom generation from HR atoms. The proposed method takes the sub-pixel displacements into account for the degradation process of HR images. Another important contribution is that it can handle a variable number of LR observations. Among a set of observed LR images, the number of LR images actually used for SR is automatically determined in each patch. There would be patches easy to estimate

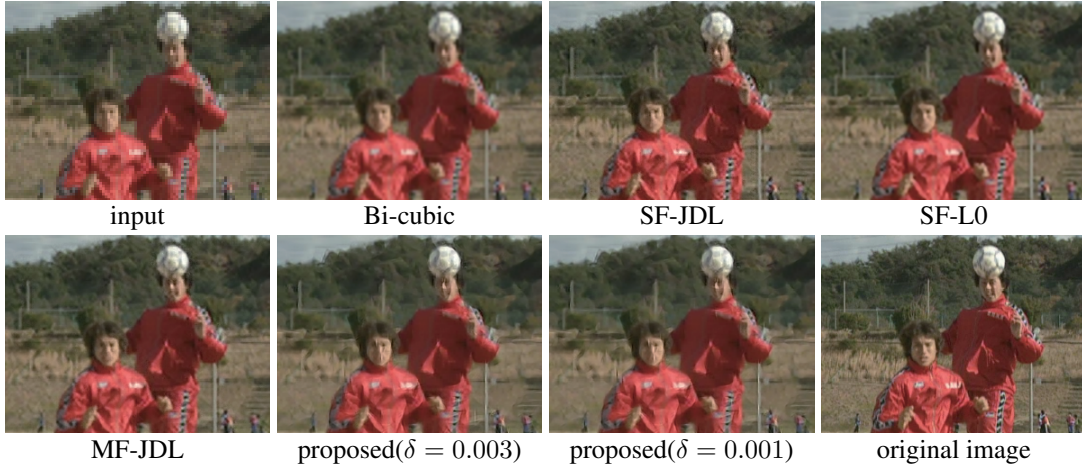


Figure 10: Images estimated from LR observations (Heading)

the displacement from the target patch, and difficult to estimate the displacement. When an LR image is composed of those two different kinds of patches, this property of automatic patch selection is useful because we can effectively utilize a subset of LR images where displacements are estimated with high confidences. When we deal with movies, it is difficult to estimate relative displacements for the regions where objects move quickly, and it is easy to estimate the relative displacements for the regions without such objects. In general, for quality of a movie, the sharpness of objects at rest or slowly moving are significant, and this property is advantageous to the proposed method.

In our future work, we will investigate an SR method for movies based on sparse coding. The proposed method is ready to be applied for movie SR by applying frame by frame, though, the current method does not consider the continuity between frames which should not be ignored for movie SR.

In example-based sparse coding SR, the sparse coding procedure is applied to certain kinds of features. Particularly, to ensure that the computed coefficient fits the most relevant part of the LR image, most of recent sparse coding based SR methods adopt high-pass filters on both y and $D_l \alpha$. The proposed method does not resort to feature extraction because of possible increase of computational costs. Finally, we are also pursuing a method to automatically determine the threshold δ using the observed LR images.

References

- [1] S. Borman and R. Stevenson, “Spatial Resolution Enhancement of Low-Resolution Image Sequences - A Comprehensive Review with Directions for Future Research,” Dept. Elect. Eng., Univ. Notre Dame, Notre Dame, in Tech. Rep., 1998.
- [2] J. Tian and K-K. Ma, “A survey on super-resolution imaging,” *Signal, Image and Video Processing*, **5**(3), pp 329–342, 2011.
- [3] R. C. Hardie, K. J. Barnard and E. E. Armstrong, “Joint MAP Registration and high resolution image estimation using a sequence of undersampled images,” *IEEE Trans. Image Process.*, **6**(12), pp. 1621-1633, 1997.

- [4] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar, “Fast and robust multiframe super-resolution,” *IEEE Transactions on Image Processing*, vol. 13, pp. 1327-1344, 2004.
- [5] A. Kanemura, S. Maeda, and S. Ishii, “Superresolution with compound Markov random fields via the variational EM algorithm,” *Neural Networks*, **22**(7), pp. 1025–1034, 2009.
- [6] W. T. Freeman, T. R. Jones, and E. C. Pasztor, “Example-based super-resolution” *IEEE Comput. Graph. Appl.*, pp. 56-65, 2002.
- [7] H. Chang, D. Yeung, and Y. Xiong, “ Super-resolution through neighbor embedding,” (CVPR), 2004.
- [8] K. Kamimura, N. Tsumura, T. Nakaguchi, Y. Miyake, and H. Motomura, “Video super-resolution using texton substitution.” In *SIGGRAPH’07: ACM SIGGRAPH 2007 posters*, page 63, New York, NY, USA, 2007. ACM.
- [9] J. Sun, N.-N. Zheng, H.Tao, and H.Y. Shum, “Image hallucination with primal sketch priors,” *Proc. 2003 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR)*, vol.2, pp.279-736, Madison, USA, June 2003.
- [10] T. Katsuki, A. Torii and M. Inoue, “Posterior-Mean Super-Resolution With a Causal Gaussian Markov Random Field Prior,” *IEEE Transactions on Image Processing*, **21**(7), pp. 3182–3193, 2012.
- [11] S. T. Roweis and L. K. Saul, “Nonlinear dimensionality reduction by locally linear embedding,” *Science*, **290**, pp. 2323–2326, 2000.
- [12] J. Yang, J. Wright, T. Huang, and Y. Ma, “Image super-resolution as sparse representation of raw image patches” *IEEE Computer Vision and Pattern Recognition (CVPR)*, 2008.
- [13] J. Yang, J. Wright, T. Huang, and Y. Ma, “Image super-resolution via sparse representation.” *IEEE Transactions on Image Processing*, 2010.
- [14] D. Barreto, L. D. Alvarez, J. Abad, “Motion estimation techniques in super-resolution image reconstruction: a performance evaluation,” In: Tsvetkov, M., Golev, V., Murtagh, F., Molina, R. (eds.) *Virtual Observatory. Plate Content Digitization, Archive Mining and Image Sequence Processing*, pp. 254–268, 2005.
- [15] G. Callico, S. Lopez, O. Sosa, J. F. Lopez, R. Sarmiento, “Analysis of fast block matching motion estimation algorithms for video super-resolution systems,” *IEEE Trans. Consum. Electron.* **54**, pp. 1430–1438, 2008.
- [16] J. K. Aggarwal and N. Nandhakumar, “On the computation of motion from sequences of images - A Review,” *Proc. IEEE*, **76**(8) pp. 917–935, 1988.
- [17] R. Szelisky and D. Scharstein, “Symmetric subpixel stereo matching,” *7th European Conference on Computer Vision (ECCV 2002)*, **2**, pp. 525–540, 2002.
- [18] Q. Tian and M. N. Huhns, “Algorithms for subpixel registration,” *Computer Vision, Graphics, and Image Processing*, **35**, pp. 220–223, 1986.
- [19] J. W. Brandt, “Improved Accuracy in Gradient-Based Optical Flow Estimation,” *Int. J. Comput. Vis.*, **25**(1), pp. 5–22, 1997.

- [20] J. L. Barron, D. J. Fleet, S. S. Beauchemin and T. A. Burkitt, "Performance of optical flow techniques," Technical Report TR-299, Dept. of Computer Science, Univ. of Western Ontario, 1992.
- [21] V. N. Dvorchenko, "Bounds on (deterministic) correlation functions with applications to registration," *IEEE Trans. Pattern Anal. Mach. Intell.*, **5**(2), pp. 206–213, 1983.
- [22] M. Shimizu and M. Okutomi, "Multi-Parameter Simultaneous Estimation on Area-Based Matching," *International Journal of Computer Vision*, **67**(3), pp. 327–342, 2006.
- [23] B. A. Olshausen and D. J. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature*, pp. 607-609, 1996.
- [24] B. A. Olshausen and D. J. Field: "How Close Are We to Understanding V1?", *Neural Computation*, **17**, pp. 1665–1699, 2005.
- [25] B. K. Natarajan: "Sparse approximate solutions to linear systems", *SIAM J. Comput.*, **24**(2), pp. 227–234, 1995.
- [26] M. Elad: "Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing", Springer-Verlag, 2010.
- [27] G. Davis, S. Mallat and M. Avellaneda, "Adaptive greedy approximations", *Constructive Approximation*, **13**(1), pp. 57–98, 1997.
- [28] S. Mallat and Z. Zhang, "Matching pursuit with time-frequency dictionaries," *IEEE Trans. Signal Processing*, **41**(12), pp. 3397–3415, 1993.
- [29] S. Chen, S. A. Billings, and W. Luo, "Orthogonal least squares methods and their application to non-linear system identification," *Int. J. Contr.*, **50**(5), pp. 1873-96, 1989.
- [30] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Proc. 27th Annu. Asilomar Conf. Signals, Systems and Computers*, 1993.
- [31] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal of Science and Computing*, **20**, pp. 33–61, 1999.
- [32] R. Tibshirani. "Regression shrinkage and selection via the Lasso," *J. Royal Statist. Soc B.*, **58**(1), pp. 267-288, 1996.
- [33] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, "Least angle regression," *Ann. Statist.*, **32**(2), pp. 407-499, 2004.
- [34] K. Engan, S. O. Aase, and J. H. Hakon, "Method of optimal directions for frame design," in *Proc. ICASSP*, **5**, pp. 2443–2446, 1999.
- [35] M. Aharon, M. Elad, and A.M. Bruckstein, "The K-SVD: An algorithm for designing of overcomplete dictionaries for sparse representation," *IEEE Trans. on Signal Processing*, **54**(11):4311-4322, 2006.
- [36] H. Lee, A. Battle, R. Raina, and A. Y. Ng, "Efficient sparse coding algorithms," in *Advances in Neural Information Processing Systems*, pp. 801-808, 2007

- [37] M. Irani, and S. Peleg, “Motion Analysis for Image Enhancement: Resolution, Occlusion, and Transparency,” *Journal of Visual Communication and Image Representation*, Vol. 4, pp. 324–335 (1993).
- [38] Y. Ueda, T. Ohta, A. Iwase, M. Seki, and N. Murata, “Super-Resolution based on Sparse Coding,” *Computational Algebraic Statistics, Theories and Applications*, pp. 10–11, 2008.
- [39] P. Wang, X. Hu, B. Xuan, J. Mu, and S. Peng, “Super Resolution Reconstruction via Multiple Frames Joint Learning,” *IEEE Conf. Multimedia and Signal Proc (CMSP)*, 2011.
- [40] A. Buades, B. Coll, and J. M. Morel, “Nonlocal image and movie denoising,” *International Journal of Computer Vision*, **76**(2), pp. 123–139, 2008.
- [41] R. Zeyde, M. Elad and M. Protter, “On single image scale-up using sparse representation,” *Proceedings of the 7th international conference on Curves and Surfaces*, pp. 711–730, 2010.
- [42] R. Rubinstein, M. Zibulevsky and M. Elad, “Double sparsity: Learning Sparse Dictionaries for Sparse Signal Approximation,” *IEEE Trans. Signal Processing*, **58**(3), pp. 1553–1564, 2010.
- [43] L. Min, L. ShiHua, W. Fu, L. Xiang, J. Hong, and J. LianJun, “Super-Resolution based on Improved Sparse Coding,” *Computer Application and System Modeling (ICASM)*, 2010.
- [44] W. Dong, L. Zhang, G. Shi, and X. Wu, “Image Deblurring and Super-Resolution by Adaptive Sparse Domain Selection and Adaptive Regularization,” *IEEE Transactions on Image Processing*, **20**(7), 2011.
- [45] X. Lu, H. Yuan, P. Yan, Y. Yuan, and X. Li, “Geometry Constrained Sparse Coding for Single Image Super-resolution,” *IEEE Computer Vision and Pattern Recognition (CVPR)*, 2012.
- [46] K-T. Chang, “Computation for Bilinear Interpolation,” *Introduction to Geographic Information Systems* (5th ed.), New York, McGraw-Hill, 2009.